



# Tecnológico de Monterrey

Inteligencia artificial avanzada para la ciencia de datos I

**Módulo 2**  
Grupo 101

**Roberto Valdez Jasso**

**A01746863**

**Módulo 2 Implementación de una técnica de aprendizaje máquina  
sin el uso de un framework**

**Estado de México**

**Martes 05 de septiembre de 2022.**

## Descripción:

Implementación de una técnica de aprendizaje maquina sin el uso de un framework.

- Crea un repositorio de GitHub para este proyecto.
- Programa uno de los algoritmos vistos en el módulo (o que tu profesor de módulo autorice) sin usar ninguna biblioteca o framework de aprendizaje máquina, ni de estadística avanzada. Lo que se busca es que implementes manualmente el algoritmo, no que importes un algoritmo ya implementado.
- Prueba tu implementación con un conjunto de datos y realiza algunas predicciones. Las predicciones las puedes correr en consola o las puedes implementar con una interfaz gráfica apoyándote en los vistos en otros módulos.
- Tu implementación debe de poder correr por separado solamente con un compilador, no debe depender de un IDE o de un “notebook”. Por ejemplo, si programas en Python, su implementación final se espera que esté en un archivo .py no en un Jupyter Notebook.
- Después de la entrega intermedia se te darán correcciones que puedes incluir en tu entrega final.

## Actividad:

Tras haber realizado la actividad, puedo denotar lo siguiente:

En primera instancia puedo comentar que momento de generar el código de modelo de IA, llegué a entender y comprender la base de como funciona, es decir, el modelo de regresión logística realizada en Python 3.9 / Pycharm, entendí el cómo las funciones mínimas que tiene que tener para la realización del mismo, ya sea el manejo de los hiperparametros como el ratio de aprendizaje en la que este va a estar aprendiendo en el set de entrenamiento, como las épocas la cual va estar ciclando para su realización, las funciones Fit que nos apoya a generar el modelo de entrenamiento con el set de datos, el calculo de pesos que nos apoya en a calcular el gradiente descendiente, y finalmente el predict un base al sigmoide podemos realizar una predicción en base a los rango de error prediseñados en la función. Ahora bien, una vez corriendo este modelo a la par con el modelo de regresión logística para la siguiente actividad y asi tambien compáralo con mi modelo generado, salieron los siguientes resultados:

## Primera Prueba

### Con Librería

```
Reporte de clasificacion :
```

	precision	recall	f1-score	support
0.0	0.82	0.88	0.85	109
1.0	0.79	0.70	0.74	69
accuracy			0.81	178
macro avg	0.80	0.79	0.79	178
weighted avg	0.81	0.81	0.81	178

### Con Modelo Manual

```
Reporte de clasificacion sin crossvalidation :
```

	precision	recall	f1-score	support
0.0	0.64	0.99	0.77	109
1.0	0.88	0.10	0.18	69
accuracy			0.65	178
macro avg	0.76	0.55	0.48	178
weighted avg	0.73	0.65	0.54	178

Con estos resultados a la mano podemos decir lo siguiente:

El realizado con la librería tiene un mayor puntaje de precisión (81%) debido a sus constante conteo y revisión del error, lo cual hace que sus épocas sen considerablemente más exactas para asi satisfacer la revisión del iteración, haciendo mas preciso el proceso con los pasos realizados con sus ratio de aprendizaje que se va modificando de acuerdo al dataframe seleccionado o tambien por la interacción de los mismos.

Ahora bien, en el modelo manual generado, podemos ver que tiene una predicción del 65% la cual no está mal pero podría ser mejor, esto se debe que el modelo a los siguientes aspectos:

- El modelo tiene un margen de error de predicción grande y no se modifica con cada interacción, lo cual gran parte de los datos pasan dentro del mismo margen
- Debido a que no se tiene un margen de error “manual” y que no se modifica por cada interacción por época, esto genera que el ratio y las épocas mejoren el modelo por cada época realizada
- Las épocas y el ratio de aprendizaje se genera en base al usuario, es decir, manualmente se agregan los valores esperados para estas variables, haciendo que el modelo y la precisión resultante se vean afectadas por lo mismo.

Por estas causas, el modelo no está “efectivo” o similar que el modelo generado por la librería, ahora bien esto se puede mejorar, si tomamos en cuenta lo anterior, podemos hacer los siguiente para mejorar el modelo:

- Que el usuario únicamente tenga la opción de agregar manualmente el ratio de aprendizaje
- Tener las épocas calculadas en base la iteración de los datos con el margen de error de estos.
- Tener un margen de error mínimo, para realizar los procesos de los datos más precisos.

- Que las épocas vayan incrementadas de acuerdo a la satisfacción de la condición al de margen de error y así obtener la mejor precisión posible de los datos.

Entre esas soluciones y otras que el momento no veo, se podría mejorar el modelo manual y aumentar su precisión y optimización , cercándolo o hacerlo similar al modelo generado por el modelo.